

Weekly Report

Yuxin Ma

11.21.2016 - 11.27.2016

Projects

Deep Learning for Visualization

- **Data Labeling** This week I and Gao Xiang were focusing on labeling scatterplots. At first we decided to use the labels in Figure 1.

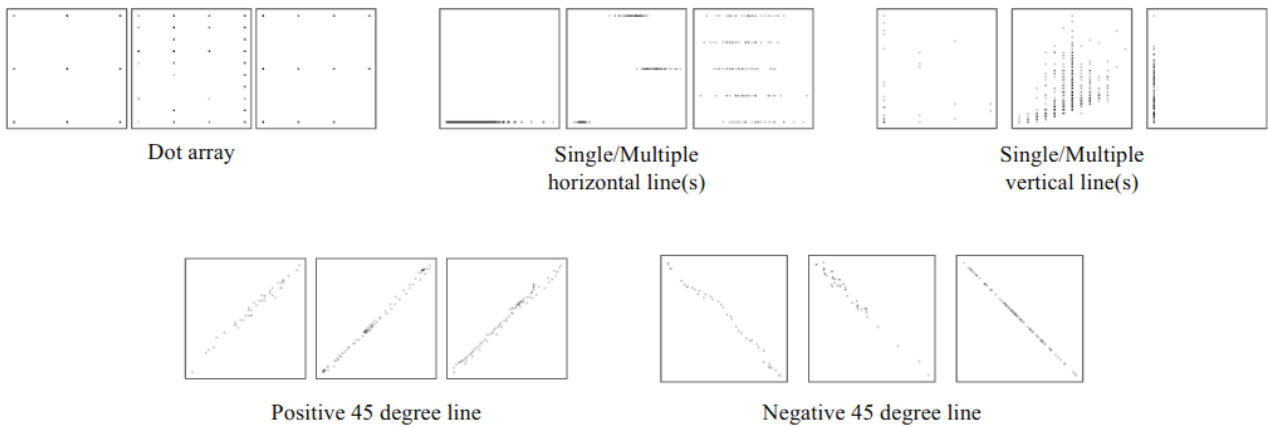


Figure 1: Five classes of scatterplot images that are easy to be detected by programs.

We have tested a series of automatic labeling strategy, i.e. detecting the shape and the data distribution in a single scatterplot to guess the actual shape in the scatterplot. However the effect is quite unusable. After discussion with Gao Xiang we decided to use the automatic labeling output as a pre-processing stage and label them by ourselves. We calculated that the work can be done in about a week by 3 people.

- **Python Programs for the Entire Workflow** This week I have written and organized the python scripts from generating scatterplot data to training a model. The automatic workflow is described in Figure 2 with a series of Python scripts which facilitates further increamental dataset supply, image generation, label pre-processing and model training. This part might be added in the paper as a background description in the future.

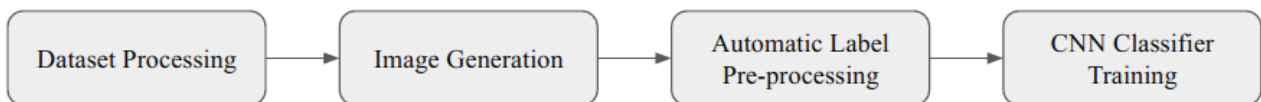


Figure 2: Organized workflow.

Plan for the Next Week

- Propose a labeling basis and label them by ourselves.
- Train the CNN classifier on the labeled result.